Updating dynamic origin-destination matrices using observed link travel speed by probe vehicles

Toshiyuki Yamamoto, Tomio Miwa, Tomonori Takeshita and Takayuki Morikawa

Nagoya Univ.

Background

P-DRGS (Probe-based dynamic route guidance system) project

Probe data collection

Calculation of traffic condition

Traffic information provision

Any other utilization of probe vehicle data?

Background

Time-dependent dynamic O-D trip matrix

- fundamental input for advanced traffic management
- difficult to observe directly

Probe vehicles (fleet cars)

- one of realized ITS technologies
- vehicle with GPS as moving sensor

Probe data have a great potential to improve the quality of dynamic O-D estimation

Existing studies

Probe information used by existing studies

- Trip origin and destination, link flow, link choice proportion, turning fraction at intersection
- Random sampling from the population

Probe vehicles in real world

- Commercial vehicles such as taxies or freight vehicles
- Probe data become seriously biased

Our approach

Even from biased sample, link travel speed/time can be useful

Two-step O-D estimation with probe data

1. Estimate population link flow from observed link travel speed of probe vehicles
2. Estimate dynamic O-D demand from estimated link flow

Advantage of two-step approach

Two-step O-D estimation with probe data

1. Estimate population link flow from observed link travel speed of probe vehicles
2. Estimate dynamic O-D demand from estimated link flow

- Flexible in applying dynamic O-D estimator in the second step
- Avoid developing time-consuming realistic micro-simulator required for one step estimation to output travel speed of each probe vehicle

Reliability of the link flow estimation

Variability of the reported link travel speed

- Biased sampling from population
- Random nature in the number of vehicles observed at each link
- Heterogeneity among drivers in driving behavior

Reliability of estimated link flow varies among links

1. Link flow estimator should provide variance of the estimates as well as point estimates for each link
2. Dynamic O-D estimator should be able to incorporate the difference in the reliability among links

Step 1: Link flow estimation

- Link performance function (Gazis et al., 1961)

$$v_a = v_{fa} \exp\left\{ -\alpha_a \left(k_a / C_a \right)^{l_a - 1} \right\}$$

v_a: average link travel speed, k_a: flow density, C_a: capacity

- Observed travel speed of probe vehicle

$$v^i_{a,t} = v^*_{a,t} + \varepsilon^i_{a,t}$$

$v^i_{a,t}$: observed travel speed, $\varepsilon^i_{a,t}$: random error

$\boxed{k_a \text{ is estimated by using Bayesian inference}}$

Bayesian inference

- Method 1: Bayesian inference assuming constant variance of the error ε

$$v_1 = \frac{\sigma_0^{-2} \cdot v_0 + n \cdot s^{-2} \cdot \bar{v}}{\sigma_0^{-2} + n \cdot s^{-2}} \qquad \sigma_1^{-2} = \sigma_0^{-2} + n \cdot s^{-2}$$

v_1, σ_1: posterior mean and standard error of v_0, σ_0: prior mean and standard error, n: sample size, s: standard error of ε

- Method 2: Bayesian inference assuming varying variance of the error across travel speed
  - Open form needs numerical integration
- Method 3: Non-Bayesian simply using sample average as point estimate and the sample size as reliability

Step 2: Dynamic O-D estimation

- The entropy maximization model (Willumsen, 1980)

$$\max Z = -\sum_{w,h_r}\left[d_w(h_r)\left\{ \ln\left(\frac{d_w(h_r)}{\hat{d}_w(h_r)} \right) - 1 \right\} \right] - \sum_{a,h} \gamma_{a,h}\left[q_a(h)\left\{ \ln\left(\frac{q_a(h)}{\hat{q}_a(h)} \right) - 1 \right\} \right]$$

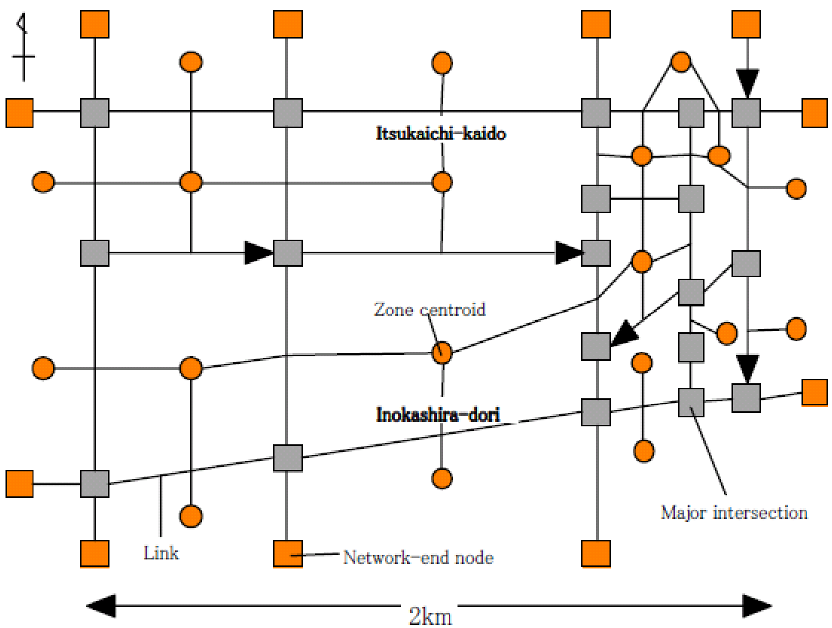
s.t. $$q_a(h) = \sum_{w,h_r} p_{aw}(h_r, h) d_w(h_r) \quad \forall a, h$$

- $\gamma_{a,h}$ is modified as dependent on the reliability of link flow estimates (conventionally set as 1)

$$\gamma_{a,h} = (1 + \beta\sigma_{a,h})^{-1}$$

Case study

- Kichijoji benchmark data set (Horiguchi et al., 1998)
  - Dynamic O-D demand and link flows are observed



- 20% of vehicles are treated as hypothetical probe cars
- Random errors are added to the true OD in order to obtain prior OD demand
- Prior distribution of v is obtained from outside of study area (Nagoya)

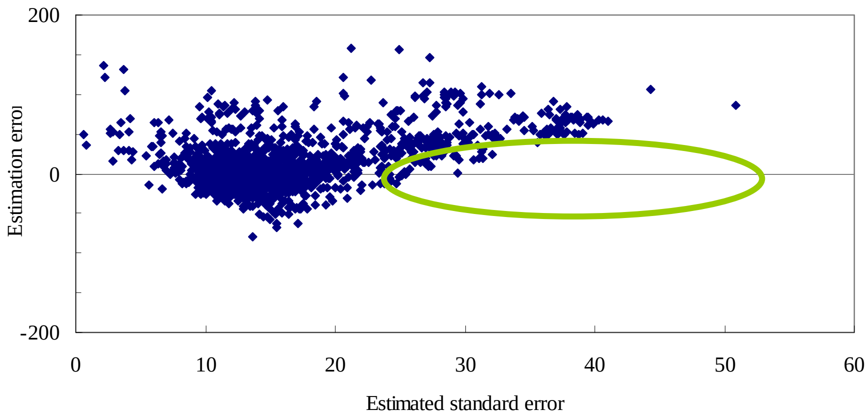
Step 1: Accuracy of the estimated link flow to be used as input to O-D estimation

	Method 1: Constant variance	Method 2: Varying variance	Method 3: Non-Byes
Corr. coef.	0.486	0.507	**0.575**
RMSE	34.80	31.80	**30.53**

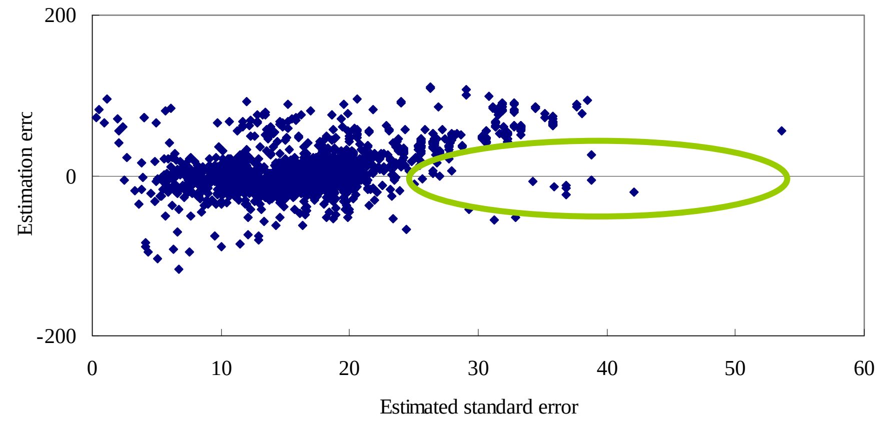
- Non-Bayesian method has the highest accuracy, implying the prior information is not effective in this case study
- However…

Distribution of estimation error

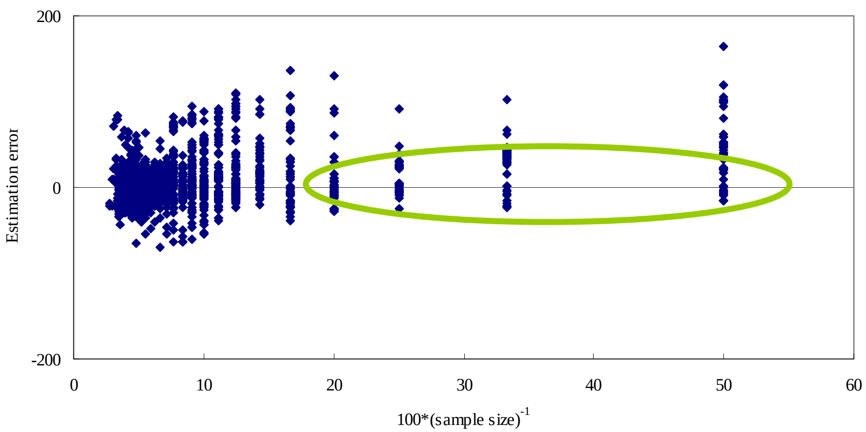
Method 1: Constant variance



Method 2: Varying variance



Method 3: Non-Bayesian



Bayesian methods provide smaller estimated standard error when the error is small

Bayesian methods correctly estimate the reliability of the link flow estimates

Step 2: Accuracy of the estimated dynamic O-D demand

	Method 1: Constant variance	Method 2: Varying variance	Method 3: Non-Byes
Corr. coef.	0.954	0.949	0.953
RMSE	1.598	1.655	1.590

If reliability of the estimated link flow is not considered ($\gamma_{a,h}$ = 1)

Corr. coef.	0.921	0.914	0.921
RMSE	2.075	2.122	2.059

Accuracy of the estimation increases by incorporating the reliability of the estimated link flow

Step 2: Accuracy of calculated link flow as output from O-D estimation

	Method 1: Constant variance	Method 2: Varying variance	Method 3: Non-Byes
Corr. coef.	0.974	0.973	0.971
RMSE	7.952	8.143	8.479

If reliability of the estimated link flow is not considered ($\gamma_{a,h}$ = 1)

Corr. coef.	0.933	0.941	0.933
RMSE	12.96	11.95	12.71

Accuracy of the estimation increases by incorporating the reliability of the estimated link flow

Conclusions

- Dynamic O-D demand estimation method using observed link travel speed is developed
  - Difference in reliability of the link flow estimates among links is explicitly incorporated

- Results of the case study suggest the improvement of accuracy by taking into account the difference in the reliability

- Advantage of Bayesian inference approach is not confirmed in this case study
  - Future research with more complete data set, probably simulation data